

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
8. März 2001 (08.03.2001)

PCT

(10) Internationale Veröffentlichungsnummer
WO 01/16327 A2

(51) Internationale Patentklassifikation⁷: C12N 15/31, C07K 14/35

(21) Internationales Aktenzeichen: PCT/DE00/02924

(22) Internationales Anmeldedatum:
28. August 2000 (28.08.2000)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
199 41 416.5 31. August 1999 (31.08.1999) DE
199 43 520.0 11. September 1999 (11.09.1999) DE

(71) Anmelder und
(72) Erfinder: NIEDERWEIS, Michael [DE/DE]; Killinger
Strasse 108, 91056 Erlangen (DE). BOSSMANN, Stefan
[DE/DE]; Haid- und Neustrasse 6, 76131 Karlsruhe (DE).

(74) Anwalt: GASSNER, Wolfgang; Nägelsbachstrasse 49A,
91052 Erlangen (DE).

(81) Bestimmungsstaaten (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Bestimmungsstaaten (*regional*): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— Ohne internationalen Recherchenbericht und erneut zu
veröffentlichen nach Erhalt des Berichts.

Zur Erklärung der Zweibuchstaben-Codes, und der anderen
Abkürzungen wird auf die Erklärungen ("Guidance Notes on
Codes and Abbreviations") am Anfang jeder regulären Ausgabe
der PCT-Gazette verwiesen.

(54) Title: METHOD FOR PRODUCING A CHANNEL-FORMING PROTEIN

(54) Bezeichnung: VERFAHREN ZUR HERSTELLUNG EINES KANALBILDENDEN PROTEINS

(57) Abstract: The invention relates to a method for producing a channel-forming protein that is present in gram-positive bacteria. The channel-forming protein is obtained by a) heterologue overexpression or b) purifying mycobacteria, whereby extraction temperature amounts to more than 50 °C.

(57) Zusammenfassung: Die Erfindung betrifft ein Verfahren zur Herstellung eines in Gram-positiven Bakterien vorkommenden kanalbildenden Proteins, wobei das kanalbildende Protein gewonnen wird durch a) heterologe Überexpression oder b) Aufreinigung aus Mycobakterien, wobei die Extraktionstemperatur mehr als 50°C beträgt.

WO 01/16327 A2

This Page Blank (uspto)

Verfahren zur Herstellung eines kanalbildenden Proteins

Die Erfindung betrifft ein Verfahren zur Herstellung eines kanalbildenden Proteins, ein kanalbildendes Protein, ein Gen
5 codierend für ein solches Protein und ein mutiertes mspA-, mspC- oder mspD- Gen, einen Plasmidvektor und ein Überexpressionssystem.

Die Erfindung betrifft allgemein das technische Gebiet der
10 Herstellung von Nanostrukturen. Zu den bisher am besten charakterisierten Nanostrukturen gehören Kohlenstoff-Nanokanäle (Yakobson, B. I. und Smalley, R. E. Fullerene nanotubes: C_{1,000,000} and beyond. *Am Sci* **85**, 324, 1997). Mit Kohlenstoff-Nanokanälen konnte gezeigt werden, daß die elek-
15 tronischen Eigenschaften durch ihre strukturellen Details kontrolliert werden. Die Synthese von Kohlenstoff-Nanokanälen erfolgt durch verschiedene Varianten von CVD (chemical vapor deposition) (Fan, S., Chapline, M. G., Franklin, N. R., Tom-
bler, T. W., Cassell, A. M. und Dai, H. Self-oriented regular
20 arrays of carbon nanotubes and their field emission properties. *Science* **283**, 512-4, 1999) und ist damit sehr aufwendig.

Aus Johnson, S. A., Ollivier, P. J. und Mallouk, T. E. Ordered mesoporous polymers of tunable pore size from colloidal
25 silica templates. *Science* **283**, 963-965 (1999) ist ein Verfahren zur Herstellung von organischen Nanokanälen auf der Grundlage eines Templates bekannt. Damit können Nanokanäle mit einem Durchmesser von 5 bis 35 nm hergestellt werden.

30 Mycobakterien gehören zu einer Untergruppe von Gram-positiven Bakterien, die Mycolsäuren besitzen und die die Gattungen *Corynebacterium*, *Nocardia*, *Rhodococcus*, *Gordona*, *Tsukamurella*, *Dietzia* einschließen.

Trias, J. und Benz, R. Permeability of the cell wall of *Mycobacterium smegmatis*. *Mol Microbiol* **14**, 283-290 (1994) beschreiben kanalbildende Proteine, nämlich Porine, in der Mycolsäure-Schicht von Mycobakterien. Biochemische oder molekulargenetische Daten über diese Porine wurden bisher nicht veröffentlicht.

Aus Lichtinger, T., Burkovski, A., Niederweis, M., Kramer, R. und Benz, R. Biochemical and biophysical characterization of the cell wall porin of *Corynebacterium glutamicum*: the channel is formed by a low molecular mass polypeptide. *Biochemistry* **37**, 15024-32 (1998) ist ein Verfahren zur Präparation von Porinen aus Corynebakterien bekannt. Dieses Verfahren ist relativ ineffizient.

Mukhopadhyay, S., Basu, D. und Chakrabarti, P. Characterization of a porin from *Mycobacterium smegmatis*. *J Bacteriol* **179**, 6205-6207 (1997) beschreiben die Extraktion von Porinen aus *M. smegmatis* mit einem Puffer mit 1 % Zwittergent durch Inkubation bei Raumtemperatur für eine Stunde. Die Ausbeuten waren schlecht und die Verunreinigung mit anderen Proteinen groß.

Aus Harth, G. et al., High-level heterologous expression and secretion in rapidly growing nonpathogenic mycobacteria of four major *Mycobacterium tuberculosis* extracellular proteins considered to be leading vaccine candidates and drug targets. *Infection and Immunity* **65**, 2321-2328 (1997) ist es bekannt, daß eine starke Expression von gering konservierten, *Mycobacterium*-spezifischen Proteinen in *E. coli* nicht möglich zu sein scheint.

Aus Senaratne, R.H. et al., Expression of a Gene for a Porin-Like Protein of the OmpA Family from *Mycobacterium tuberculosis* H37Rv. *J Bacteriol* **180**, 3541-3547 (1998) ist die Expression eines Gens für ein porinartiges Protein aus *Mycobacterium tuberculosis* H37Rv in *E. coli* bekannt. Die Expression des Gens verursacht einen Abbruch des Bakterienwachstums, weil das exprimierte Protein offensichtlich toxisch für *E. coli* ist. Es können lediglich geringe Mengen des Proteins aus *E. coli* kurz vor dem Absterben der Zellen isoliert werden.

10

Aufgabe der Erfindung ist es, ein verbessertes Verfahren zur Herstellung eines kanalbildenden Proteins anzugeben.

Diese Aufgabe wird durch die Merkmale der Ansprüche 1, 26, 27, 28, 30, 32, 36, 37, 38, 40 und 41 gelöst. Zweckmäßige Weiterbildungen ergeben sich aus den Merkmalen der Ansprüche 2 bis 25, 29, 31, 33, 34, 35, 39 sowie ggf. 37 und 38.

Nach Maßgabe der Erfindung ist ein Verfahren zur Herstellung eines in Gram-positiven Bakterien vorkommenden kanalbildenden Proteins vorgesehen, wobei das kanalbildende Protein gewonnen wird durch

a) heterologe Überexpression oder

25

b) Aufreinigung aus *Mycobakterien*, wobei die Extraktions-temperatur mehr als 50°C beträgt.

Unter den kanalbildenden Proteinen werden Proteine verstanden, welche einen wassergefüllten Kanal oder eine wassergefüllte kanalartige Struktur bilden können. Solche Proteine kommen natürlicherweise, insbesondere in der Zellwand von Bakterien, vor. Sie können kanalartige Strukturen oder Kanäle

30

mit einem Durchmesser bis zu 3 nm und sogar darüber bilden. Die Länge kann bis zu 10 nm und mehr betragen. Die kanalartigen Strukturen oder Kanäle können aus mehreren, insbesondere vier oder acht, Untereinheiten gebildet sein.

5

Das erfindungsgemäße Verfahren ist wesentlich effizienter als die bisher beschriebenen Verfahren, bietet die Möglichkeit einer weitgehenden Automatisierung der chromatographischen Aufreinigung und ermöglicht eine drastisch erhöhte Ausbeute.

10

Das Gram-positive Bakterium kann ein mindestens eine Mycol-säure enthaltendes Bakterium sein. Nach einer Ausgestaltung ist das Bakterium ein Mykobakterium, vorzugsweise *Mycobacterium smegmatis*.

15

Das kanalbildende Protein kann ein Porin sein. Bevorzugt wird ein Porin, das im wesentlichen gegenüber organischen Lösungsmitteln chemisch stabil und/oder bis zu einer Temperatur von 80°C, vorzugsweise 100°C, thermisch stabil ist. Stabil bedeutet, daß die kanalartige Struktur des Proteins erhalten bleibt und im wesentlichen keine Denaturierung des Proteins stattfindet.

20

Das Porin ist vorzugsweise MspA, MspC, MspD, ein Fragment dieser Porine, ein zu diesen Porinen oder deren Fragmenten homologes Protein oder ein von einer Sequenz dieser Porine abgeleitetes Protein ist. MspA weist die Sequenz der Aminosäuren 28 - 211 der Sequenz 3 (siehe unten), MspC die Sequenz 7 und MspD die Sequenz 9 auf. Das zu den genannten Porinen oder deren Fragmenten homologe Protein weist eine ähnliche Struktur wie diese Porine oder Fragmente auf. Mindestens 20 % der Aminosäuren sind identisch oder homolog zu den Aminosäuren dieser Porine oder Fragmente. Eine Aminosäure in

25

30

einem Protein ist zu einer anderen Aminosäure homolog, wenn sie durch die andere Aminosäure ersetzt werden kann, ohne daß dadurch die Funktion oder Struktur des Proteins wesentlich beeinflusst wird. Bei dem von einer der Sequenzen der Porine abgeleiteten Protein können gegenüber diesen Sequenzen einzelne oder mehrere Aminosäuren fehlen oder durch andere Aminosäuren oder Aminosäureanaloga ersetzt sein.

Die genannten Proteine eignen sich wegen ihrer überraschend hohen chemischen und thermischen Stabilität besonders gut zur Herstellung von Nanostrukturen.

Eine gute Ausbeute wird erzielt, wenn die heterologe Überexpression in *E. coli* oder Mycobakterien durchgeführt wird. Zweckmäßigerweise wird zur Überexpression ein für ein kanalbildendes Protein, vorzugsweise ein Porin, codierendes Gen benutzt. Vorteilhaft ist es weiter, daß zur Überexpression ein mspA-Gen gemäß Sequenz 1 (siehe unten), ein mspC-Gen gemäß Sequenz 6 oder ein mspD-Gen gemäß Sequenz 8 benutzt wird. Zur Überexpression kann auch ein von den Sequenzen 1, 6 oder 8 abgeleitetes mutiertes Gen benutzt werden, wobei die Mutation so ausgebildet ist, daß die chemische und thermische Stabilität sowie die kanalartige Struktur des exprimierten Proteins im wesentlichen denen von MspA, MspC oder MspD entspricht. Die Mutation kann auch im wesentlichen in einer Angleichung der Codons des mspA-Gens, des mspC-Gens oder des mspD-Gens an die Codons der in *E. coli* hoch exprimierten Gene bestehen. Solche Codons sind aus Nakamura, T. et al., Two types of linkage between codon usage and gene-expression levels. *FEBS Lett.* **289**, 123-125 (1991) bekannt.

Zur Überexpression kann auch ein mutiertes mspA-, mspC- oder mspD-Gen benutzt werden, wobei die Mutation im wesentlichen

darin besteht, daß der GC-Gehalt auf weniger als 66% vermindert ist.- Die Anpassung der Codon-Benutzung verbessert die Überexpression von MspA, MspC und MspD in *E. coli* erheblich.

- 5 Durch Herstellung des kanalbildenden Proteins MspA aus *E. coli* kann die Ausbeute gegenüber dem oben beschriebenen Verfahren zur Präparation des nativen Proteins noch einmal um den Faktor 10 bis 20 gesteigert werden.
- 10 Es hat sich als vorteilhaft erwiesen, zur Überexpression das synmspA-Gen gemäß Sequenz 4 zu benutzen. Dazu kann ein zur Überexpression in *E. coli* geeigneter Vektor verwendet werden, in den das synmspA-Gen gemäß Sequenz 4 eingesetzt ist. Solche geeigneten Vektoren sind z.B. von Hannig, G. und Makrides, S.C. in Trends in Biotechnology, 1998, Vol. 16, pp54 be-
- 15 beschrieben. Der Offenbarungsgehalt dieses Dokuments wird hiermit einbezogen.

- Es hat sich weiter als vorteilhaft erwiesen, die kanalbildenden Proteine mittels nicht-ionischer oder zwitterionischer Detergentien aus der Zellwand von Gram-positiven Bakterien zu gewinnen. Die Detergentien können aus der folgenden Gruppe ausgewählt sein: Isotridecylpoly(ethyleneglycolether)_n, Alkylglucoside, besonders Octylglucosid, Alkylmaltoside, besonders Dodecylmaltosid, Alkylthioglucoside, besonders Oc-
- 20 tylthioglucosid, Octyl-Polyethylenoxide und Lauryldimethylaminoxid. Es ist zweckmäßigerweise eine zweifache oder höhere kritische micellare Konzentration (CMC) in einem Phosphatpuffer (100 mM Na₂HPO₄/NaH₂PO₄, pH 6,5, 150 mM NaCl) eingestellt
- 25 worden. - Die zwitterionischen und nicht-ionischen Detergentien lösen insbesondere das kanalbildende Protein MspA sehr selektiv und mit guter Ausbeute aus der Zellwand von *M. smegmatis*.
- 30

Es hat sich weiter als zweckmäßig erwiesen, daß die Temperatur bei der Extraktion zwischen 80 und 110 °C, vorzugsweise zwischen 90 und 100 °C, und/oder die Extraktionszeit 5 bis 120 Minuten, vorzugsweise 25 - 35 Minuten, beträgt. Vorteilhaft ist weiter die Benutzung eines Puffers mit einer Ionenstärke von mehr als 50 mM NaCl oder Na-Phosphat.

Insbesondere eine Durchführung der Extraktion bei 100 °C, die Verwendung eines Puffers mit hoher Ionenstärke sowie zwittrionischer und nicht-ionischer Detergentien verbessern das Extraktionsverfahren für Porine aus *Mycobacterium smegmatis*. Es bietet gegenüber den bisherigen Verfahren zur Aufreinigung solcher Proteine mit Hilfe organischer Lösungsmittel oder der Extraktion bei Raumtemperatur folgende Vorteile:

- aa) Verzicht auf organische Lösungsmittel
- bb) geringe Verunreinigungen mit anderen Proteinen
- cc) effiziente Extraktion

Es ist auch möglich, MspA zur Aufreinigung in Dimethylsulfoxid bei einer Temperatur im Bereich von 50 - 110 °C zu lösen; danach kann die Lösung vom Rückstand getrennt und MspA durch Abkühlen ausgefällt werden.

Vorteilhafterweise wird das kanalbildende Protein zur Aufreinigung, insbesondere mittels Aceton, ausgefällt. Dabei kann es zu einer Anreicherung des kanalbildenden Proteins gegenüber nicht ausfallenden Proteinen kommen. Weiterhin ist es vorteilhaft, das Protein zur Aufreinigung einer Ionenaustauscher-Chromatographie, insbesondere einer Anionenaustauscher-Chromatographie, zu unterwerfen. Zur weiteren Aufreini-

gung kann das kanalbildende Protein einer Größenausschluß-Chromatographie unterworfen werden.

Das durch heterologe Überexpression gewonnene kanalbildende Protein kann durch Erhöhen der lokalen Konzentration des kanalbildenden Proteins renaturiert werden. Das Erhöhen der lokalen Konzentration kann durch elektrophoretische Anreicherung, insbesondere durch Anlegen einer Gleichspannung, durch Ausfällen oder durch Adsorption an einer Oberfläche, insbesondere einer Membran, erfolgen. Zweckmäßig ist das Anlegen einer Gleichspannung im Bereich von 50 V für eine Zeit von etwa 30 Minuten.

Gegenstand der Erfindung ist weiterhin ein kanalbildendes Protein aus einem Gram-positiven Bakterium hergestellt nach dem erfindungsgemäßen Verfahren.

Das Gram-positive Bakterium kann ein Mycolsäure enthaltendes Bakterium sein, wobei zweckmäßigerweise das Bakterium ein Mykobakterium, vorzugsweise *Mycobacterium smegmatis*, ist.

Von besonderem Vorteil ist es, daß das kanalbildende Protein ein Porin ist, das im wesentlichen gegenüber organischen Lösungsmitteln chemisch stabil ist. Das Porin ist vorzugsweise im wesentlichen bis zu einer Temperatur von 80°C, vorzugsweise 100°C, thermisch stabil. Es kann sich dabei um das Porin MspA, MspC, MspD, ein Fragment dieser Porine, ein zu diesen Porinen oder deren Fragmenten homologes Protein oder ein von der Sequenz dieser Porine abgeleitetes Protein handeln. Die chemische und thermische Stabilität sowie die kanalartige Struktur des abgeleiteten Proteins kann im wesentlichen derjenigen der Proteine MspA, MspC oder MspD entsprechen. Es ist aber auch denkbar, daß weitere hier nicht genannte Porine

diese Eigenschaften aufweisen und damit vom Gegenstand der vorliegenden Erfindung umfaßt sind.

Die erfindungsgemäßen kanalbildenden Proteine haben die folgenden Vorteile:

aaa) Soweit die kanalbildenden Proteine in der Zellwand von *M. smegmatis* enthalten sind, lassen sie sich daraus in organischen Lösungsmitteln (z. B. $\text{CHCl}_3/\text{MeOH}$) lösen, ohne zu denaturieren. Die Fähigkeit zur Kanalbildung bleibt in organischen Lösungsmitteln erhalten.

bbb) Sie lassen sich mit Aceton fällen, ohne zu denaturieren.

ccc) Sie überstehen selbst Kochen in Detergentien (z. B. 10 Minuten in 3% SDS), ohne zu denaturieren.

Diese extreme Stabilität der erfindungsgemäßen Proteine gegenüber chemischer und thermischer Denaturierung ermöglicht deren Verwendung zur Herstellung von technisch verwertbaren Nanostrukturen.

Nach Maßgabe der Erfindung wird weiterhin ein für ein erfindungsgemäßes kanalbildendes Protein codierendes Gen beansprucht. Das kann das mspA-Gen gemäß Sequenz 1, das mspC-Gen gemäß Sequenz 6 oder das mspD-Gen gemäß Sequenz 8 sein.

Als weiterer Gegenstand kommt auch ein mutiertes mspA-Gen, mspC-Gen oder mspD-Gen in Betracht, wobei die Mutation im wesentlichen in einer Angleichung der Codons des mspA-Gens, des mspC-Gens oder des mspD-Gens an die Codons der in *E. coli* hoch exprimierten Gene besteht. Die Mutation kann im wesentlichen darin bestehen, daß der GC-Gehalt auf weniger als 66%

vermindert ist. Das mutierte Gen kann aber auch von einer der Sequenzen 1, 6 oder 8 abgeleitet und so ausgebildet sein, daß die chemische und thermische Stabilität sowie die kanalartige Struktur des exprimierten Proteins im wesentlichen der von MspA, MspC oder MspD entspricht. Weitere hier nicht genannte Mutationen sind für den Fachmann ebenfalls denkbar. Gene, die zur Ausbildung der erfindungsgemäßen kanalartigen Proteine führen, sind vom beanspruchten Schutzzumfang umfaßt. Z.B. ein mutiertes mspA-Gen, wobei das mutierte Gen das synmspA-Gen gemäß Sequenz 4 (siehe unten) ist.

Ein weiterer Gegenstand der Erfindung ist der Plasmidvektor pMN501 und ein Überexpressionssystem, bei dem *E. coli* diesen Plasmidvektor enthält.

Nachfolgend werden anhand der Figuren Beispiele der Erfindung erläutert. Es zeigen:

Fig. 1a-c die Temperaturabhängigkeit der Extraktion von MspA aus *M. smegmatis* in gelelektrophoretischer Darstellung,

Fig. 2 die Reinigung von MspA aus *M. smegmatis* in gelelektrophoretischer Darstellung,

Fig. 3 die Reinigung von MspA aus *E. coli* in gelelektrophoretischer Darstellung,

Fig. 4 die Konstruktion des Plasmidvektors pMN501,

Fig. 5 eine schematische Ansicht einer Vorrichtung zur Renaturierung,

Fig. 6 ein renaturiertes MspA in gelelektrophoretischer Darstellung und

5 Fig. 7a-c Modifikationen des Kanalproteins MspA in elektronenmikroskopischer Darstellung.

Bei allen gelelektrophoretischen Darstellungen von Proteinen sind die Proteine 30 Minuten bei Raumtemperatur in Probenpuffer (40 mM Tris(hydroxymethyl)aminomethan, pH 7,0, 3% SDS, 8% Glycerin, 0,1% Serva Blau G) inkubiert und dann gelelektrophoretisch nach ihrer Größe aufgetrennt worden.

Die Figuren 1a-c zeigen bei verschiedenen Temperaturen aus *M. smegmatis* extrahierte Proteine. Fig. 1a zeigt ein mit Coomassie Blau gefärbtes 10%-iges SDS-Polyacrylamidgel. Spur M: Massenstandard mit 200, 116,3, 97,4, 66,3, 55,4, 36,5, 31, 21,5 und 14,4 kDa. Spuren 1 bis 8: je 12 µl von bei 30, 40, 50, 60, 70, 80, 90 oder 100°C erhaltenen Extrakten.

20 Fig. 1b zeigt einen Immunoblot eines 8%-igen SDS-Polyacrylamidgels auf einer PVDF-Membran. Die Proteine wurden unter Verwendung eines gegen MspA gerichteten Antiserums aus Kaninchen und einer Chemilumineszenz-Reaktion (ECL-Detektionssystem, Amersham-Pharmacia, Wien, Österreich) sichtbar gemacht. Spur M: Massenstandard mit 97,4, 68, 46, 31, 20,1 und 14,4 kDa. Spuren 1 bis 3: Je 2 µl eines bei 30, 40 oder 50°C erhaltenen Extrakts. Spuren 4 bis 8: Je 1 µl eines bei 60, 70, 80, 90 oder 100°C erhaltenen Extrakts. Spur 9: 1 ng MspA.

30

Fig. 1c zeigt ein mit Silber gefärbtes 8%-iges SDS-Polyacrylamidgel. Spur M: Massenstandard mit 200, 116,3, 97,4, 66,3, 55,4, 36,5, 31, 21,5 und 14,4 kDa. Spur 1 und 2:

Je 15 μ l eines bei 30 oder 40°C gewonnenen Extrakts. Spur 3: 10 μ l eines bei 50°C gewonnenen Extrakts. Spuren 4 bis 8: Je 4 μ l eines bei 60, 70, 80, 90 oder 100°C erhaltenen Extrakts. Spur 9 zeigt 270 ng aufgereinigtes MspA.

5

Fig. 2 zeigt ein mit Coomassie Blau gefärbtes 10 %iges SDS-Polyacrylamidgel. Spur M: Massenstandard: 200, 116,3, 97,4, 66,3, 55,4, 36,5, 31, 21,5 und 14,4 kDa. Spur 1: 40 μ g Protein eines aus *M. smegmatis* mit POP05-Puffer (100 mM Na₂HPO₄/NaH₂PO₄, pH 6,5, 0,1 mM EDTA, 150 mM NaCl, 0,5 % Octylpolyethylenoxid (OPOE)) gewonnenen Extrakts. Spur 2: 40 μ g Protein aus dem Extrakt nach Fällung mit Aceton. Spur 3: 4 μ g Protein aus vereinigten MspA enthaltenden Fraktionen einer Anionenaustauscher-Chromatographie. Spur 4: 4 μ g Protein der MspA enthaltenden Fraktionen nach Fällung mit Aceton. Spur 5: 4 μ g Protein aus vereinigten MspA enthaltenden Fraktionen nach Größenausschluß-Chromatographie. Die Sequenzen des mspA-Gens, des mspA-Gens + Promotor sowie des MspA-Proteins mit vermuteter Signalsequenz sind im Sequenzprotokoll als Sequenzen 1 - 3 wiedergeben.

Fig. 3 zeigt die Reinigung des Kanalproteins MspA aus *E. coli*. Die Proteine wurden mit einem 10 %igen SDS-Polyacrylamidgel nach Größe getrennt. Das Gel wurde mit Coomassie Blau gefärbt. Spur 1: Lysat von *E. coli* BL21(DE3)/pMN501 vor der Induktion durch IPTG. Spur 2: Lysat von *E. coli* BL21(DE3)/pMN501 nach der Induktion durch IPTG. Spur 3: Massenstandard: 200, 116, 97, 66, 55, 36,5, 31, 21,5, 14,4 und 6 kDa. Die Proben wurden 30 Minuten bei 37 °C inkubiert, bevor sie auf das Gel aufgetragen wurden.

In Fig. 4 ist schematisch die Konstruktion des Plasmids pMN501 zur Überexpression von MspA in *E. coli* BL21(DE3) dargestellt. Die verwendeten Abkürzungen bedeuten:

- 5 lacI: Gen codierend für den Laktose-Repressor
- nptI: Gen codierend für die Neomycinphosphotransferase; sie
 vermittelt Kanamycinresistenz
- Ori: Replikationsursprung
- RBS: Ribosomenbindestelle

10

Fig. 5 zeigt schematisch eine Vorrichtung zur Renaturierung von monomerem MspA. Eine Pipettenspitze aus Polyethylen von 5 cm Länge, deren unteres Ende nach ca. 2 mm abgeschnitten wurde, wurde mit einer 1.7 %igen Agarose-Lösung (in TAE-Puffer) gefüllt. Eine Bleistiftmine (Typ: Eberhard Faber, 3H) wurde auf eine Länge von 5 cm gekürzt. Ein Polypropylengefäß ohne Deckel wurde mit 60 µl einer Lösung mit 5 µg denaturiertem MspA gefüllt und die Pipettenspitze und die Bleistiftmine in die Lösung gestellt. Dann wurde die Pipettenspitze als Kathode und die Bleistiftmine als Anode angeschlossen.

20

Fig. 6 zeigt die Renaturierung von denaturiertem MspA. Die Proteine wurden mit einem 10 %igen SDS-Polyacrylamidgel nach Schägger getrennt (Schägger, H. and von Jagow, G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* 166, 368-79 (1987)). Das Gel wurde mit Silber gefärbt. Spur M: Massenstandard: 116, 97, 66, 55, 36.5, 31, 21.5, 14.4 kDa; Spur 1: 800 ng denaturiertes MspA. Spur 2: 800 ng MspA nach der Renaturierungsreaktion. Die Proben wurden 30 Minuten bei 37 °C inkubiert, bevor sie auf das Gel aufgetragen wurden.

30

Die Fig. 7a bis 7c zeigen elektronenmikroskopische Aufnahmen von Modifikationen des Kanalproteins MspA aus *M. smegmatis*. Die Herstellung der Proben erfolgt nach folgendem Protokoll: Ein Milliliter einer Lösung des Kanalproteins MspA ($c(\text{MspA}) = 17,2 \times 10^{-9} \text{ mol/L}$, 100 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$, pH 6,5, 150 mM NaCl, 0,10 g/L SDS) werden bei 24,5°C in einem Ultraschallbad dispergiert. Zwischen der Flüssigkeitsoberfläche und einer HOPG (Kohlenstoff) - Oberfläche ($1,0 \text{ mm}^2$) wird ein konstanter Abstand von 5,0 cm eingestellt. Die HOPG-Oberfläche wird für 20 Sekunden den dispergierten Flüssigkeitströpfchen ausgesetzt.

In Fig. 7a liegen isolierte Kanalproteine vor. In Fig. 7b ist eine Bänderstruktur erkennbar, die große Poren mit einem Durchmesser von 12 nm aufweist. Aus Fig. 7c ist ersichtlich, daß die Bänderstruktur zwei Typen von Kanälen beitzzt, nämlich erste Kanäle mit einem kleinen Durchmesser von etwa 2.4 nm und zweite Kanäle mit größeren Durchmesser von etwa 9.0 bis 10,0 nm.

20

Beispiel 1: Extraktion von MspA aus *M. smegmatis* bei verschiedenen Temperaturen

10 mg *M. smegmatis* mc²155-Zellen (Naßgewicht) wurden mit PBS (100 mM Natriumphosphat, pH 7,0, 150 mM NaCl, 0,1 mM EDTA) gewaschen und in 150 µl PG05-Puffer (0,5 % Isotridecylpolyethylenglycolether, 100 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$, 0,1 mM EDTA, 150 mM NaCl, pH 6,5) resuspendiert. Die suspendierten Zellen wurden für 30 Minuten bei 30, 40, 50, 60, 70, 80, 90 oder 100°C inkubiert. Die Proben wurden für 10 Minuten auf Eis gekühlt und für 10 Minuten bei 4°C zentrifugiert. Das Volumen des Überstands wurde durch Verdampfen von 120 auf 10 µl reduziert. Die Proteine wurden gelelektrophoretisch nach ihrer Größe getrennt, wie es in den Figuren 1a-c gezeigt ist. Der

Anteil von MspA an dem insgesamt extrahierten Protein nimmt bei Temperaturen über 50°C deutlich zu. Andere Proteine werden bei diesen Temperaturen kaum noch extrahiert.

5 Beispiel 2: Aufreinigung von MspA aus *M. smegmatis*

10 g *M. smegmatis* (Naßgewicht) wurden mit PBS gewaschen, in 35 ml POP05-Puffer resuspendiert und unter Rühren für 30 Minuten in einem Wasserbad gekocht. Die Zellsuspension wurde für 10 Minuten auf Eis gekühlt und bei 4°C für 15 Minuten bei 27000 g zentrifugiert. 42 ml des Überstands wurden vorsichtig mit einem gleichen Volumen eisgekühlten Acetons gemischt. Die Mischung wurde eine Stunde auf Eis gekühlt und bei 4°C für 15 Minuten bei 8000 g zentrifugiert. Das ausgefällte Protein wurde in 10 ml 25 mM AOP05 (*N*-(2-Hydroxyethyl)piperazin-*N'*-2-ethansulfonsäure (Hepes), pH 7,5, 10 mM NaCl, 0,5 % OPOE) gelöst und auf eine Anionenaustauscher-Säule POROS 20HQ mit einem Volumen von 1,7 ml (Perseptive Biosystems, Cambridge, MA) geladen. Nach Waschen der Säule mit 14 ml AOP05 wurden gebundene Proteine mit einem Gradienten von 100% AOP05 bis 100 % BOP05 (25 mM Hepes, pH 7,5, 2 M NaCl, 0,5 % OPOE) über 34 ml eluiert. Es wurden 90 1 ml-Faktionen gesammelt und gelelektrophoretisch analysiert. MspA eluierte zwischen 0,48 und 0,74 M NaCl mit einem Peak bei 0,57 M NaCl. Die 4 Fraktionen mit der höchsten MspA-Menge wurden vereinigt und mit Aceton gefällt. Das resultierende Pellet wurde in 600 µl AOP05 aufgenommen, auf Eis inkubiert und bei 4°C für 5 Minuten zentrifugiert um unlösliches Material zu entfernen. Die resultierende Proteinlösung wurde auf eine Superdex 200 Gelfiltrations-Säule mit einem Volumen von 24 ml (Pharmacia, Freiburg, Deutschland) aufgetragen. Proteine wurden mit 48 ml AOP05 bei einer Flußrate von 0,2 ml/Minute eluiert. 50 1 ml-Faktionen wurden gesammelt und gelelektrophoretisch analysiert. Fraktionen mit nahezu reinem MspA wurden vereinigt.

Die einzelnen Reinigungsstufen sind aus Fig. 2 ersichtlich. Die Ausbeute beträgt 700 µg. 1 µg dieser Probe zeigte in einem silbergefärbten SDS-Polyacrylamidgel keinerlei Kontaminationen mit anderen Proteinen (nicht gezeigt). MspA ist bis
5 oder nahezu bis zur Homogenität aufgereinigt worden.

Beispiel 3: Verfahren zur Präparation des Kanalproteins MspA aus *E. coli*

10 Zur weiteren Erhöhung der Ausbeute an MspA wird eine Überexpression des entsprechenden Gens vorgeschlagen. Zunächst wird das *mspA*-Gen kloniert, das für das Kanalprotein MspA aus *Mycobacterium smegmatis* mc²155 kodiert. Es wird das T7-Expressionssystem für die Überexpression des *mspA*-Gens ge-
15 wählt.

Das *mspA*-Gen wird aus dem Plasmid pPOR6 über PCR amplifiziert. In der nativen *mspA*-Sequenz werden alle Codons verändert, die in stark exprimierten Genen aus *Escherichia coli*
20 selten vorkommen. Im Sequenzprotokoll (siehe unten), sind in Sequenz 4 alle eingeführten Mutationen aufgelistet. Diese *synmspA* genannte DNA wird nach der Methode von Stemmer (Stemmer, W. P., Cramer, A., Ha, K. D., Brennan, T. M. and Heyneker, H. L. Single-step assembly of a gene and entire plasmid
25 from large numbers of oligodeoxyribonucleotides. *Gene* **164**, 49-53 (1995)) durch Assemblierung von Oligonucleotiden synthetisiert und anstelle des *mspA*-Gens in den Vektor pMN500 eingesetzt. Das resultierende Plasmid pMN501 (Fig. 4) vermittelt in Zellen von *E. coli* BL21(DE3) eine starke Expression
30 von denaturiertem MspA-Monomer (20 kDa) nach Induktion mit IPTG. Das so exprimierte rMspA genannte Protein weist die Sequenz 5 (siehe unten) auf.

Beispiel 4: Aufreinigung von MspA aus *E. coli*

Ein Liter LB-Medium mit 30 $\mu\text{g/mL}$ Kanamycin wird mit *E. coli* BL21(DE3)/pMN500 beimpft und bei 37°C bis zu einer OD₆₀₀ von 0.6 geschüttelt. Dann wird mit 1 mM IPTG induziert und die Zellen noch sechs Stunden bei 37°C bis zu einer OD₆₀₀ von 2.2 geschüttelt. Die Zellen werden in 40 mL A-Puffer (25 mM HEPES, pH 7,5, 10 mM NaCl) resuspendiert und durch zehnminütiges Kochen in Wasser aufgeschlossen. Nach einer zehnminütigen Inkubation auf Eis werden die Zelltrümmer und die ausgefallenen Proteine durch Zentrifugation bei 10000 g für 10 Minuten abgetrennt. Der Überstand wird an einem Anionenaustauscher (POROS HQ20) mit einem linearen NaCl-Gradienten von 10 mM bis 2 M NaCl getrennt. Monomeres MspA eluiert bei 350 mM NaCl. Um höhermolekulare Proteine abzutrennen, werden die Fraktionen mit MspA vereinigt und es wird eine Gelfiltration durchgeführt. Die Ausbeute beträgt 10 mg MspA mit einer Reinheit von über 95 % (Daten nicht gezeigt).

Beispiel 5: Elektrochemische Assemblierung des Kanalproteins MspA

Durch die Überexpression von MspA in *E. coli* ist es zwar leicht möglich, das Kanalprotein mit einer guten Ausbeute zu isolieren. Das gewonnene Protein liegt zum großen Teil in inaktiver Form vor. Die Überführung in die aktive Form bzw. Renaturierung von monomerem MspA kann nach folgendem Protokoll erfolgen:

Die Renaturierung kann in einer für diesen Zweck entwickelten Apparatur stattfinden (Fig. 5). Die Renaturierungsreaktion wird mit 5 μg MspA in monomerer Form in dieser Reaktionsapparatur durch Anlegen einer Spannung von 50 V für 30 Minuten durchgeführt. Zum Schluß wird die Spannung für

fünf Sekunden umgepolt, um an der Bleistiftmine adsorbiertes Porin wieder zu lösen.

Das Protein wird nach der oben beschriebenen Renaturierungsreaktion in einem Proteingel untersucht (Fig. 6). Dabei stellt sich heraus, daß ein großer Teil des Proteins zu oligomeren Einheiten assembliert ist. Durch Rekonstitutionsexperimente kann gezeigt werden, daß das MspA in dieser Form wieder hohe Kanalaktivität besitzt. Das beweist, daß die Renaturierung von MspA durch geringe Gleichspannungen möglich ist.

Diese Renaturierungsreaktion ist sehr einfach durchzuführen und ist damit ein wichtiger Bestandteil der Präparation von funktionalem Kanalprotein MspA aus überproduzierenden *E. coli*.

Liste der Sequenzen:

1. mspA-Gen, translatiert
- 20 2. mspA-Gen + Promotor, translatiert
3. MspA-Proteins mit vermuteter Signalsequenz
4. synmspA-Gen, translatiert
5. rMspA-Protein
6. mspC-Gen
- 25 7. MspC-Protein
8. mspD-Gen
9. mspD-Protein

Patentansprüche

1. Verfahren zur Herstellung eines in Gram-positiven Bakterien vorkommenden kanalbildenden Proteins, wobei das kanalbildende Protein gewonnen wird durch
 - a) heterologe Überexpression oder
 - b) Aufreinigung aus Mycobakterien, wobei die Extraktions-temperatur mehr als 50°C beträgt.
2. Verfahren nach Anspruch 1, wobei das Gram-positive Bakterium ein mindestens eine Mycolsäure enthaltendes Bakterium ist.
3. Verfahren nach einem der vorhergehenden Ansprüche, wobei das Bakterium ein Mykobakterium, vorzugsweise *Mycobacterium smegmatis*, ist.
4. Verfahren nach einem der vorhergehenden Ansprüche, wobei das kanalbildende Protein ein Porin ist.
5. Verfahren nach einem der vorhergehenden Ansprüche, wobei das Porin im wesentlichen gegenüber organischen Lösungsmitteln chemisch stabil ist.
6. Verfahren nach einem der vorhergehenden Ansprüche, wobei das Porin im wesentlichen bis zu einer Temperatur von 80°C, vorzugsweise 100°C, thermisch stabil ist.
7. Verfahren nach einem der vorhergehenden Ansprüche, wobei das Porin MspA, MspC, MspD, ein Fragment dieser Porine, ein zu diesen Porinen oder deren Fragmenten homologes Protein

oder ein von einer Sequenz dieser Porine abgeleitetes Protein ist.

8. Verfahren nach einem der vorhergehenden Ansprüche, wobei
5 die heterologe Überexpression in *E. coli* oder Mycobakterien durchgeführt wird.

9. Verfahren nach einem der vorhergehenden Ansprüche, wobei
10 zur Überexpression ein für ein kanalbildendes Protein, vorzugsweise ein Porin, codierendes Gen benutzt wird.

10. Verfahren nach einem der vorhergehenden Ansprüche, wobei
zur Überexpression ein mspA-Gen gemäß Sequenz 1, ein mspC-Gen
gemäß Sequenz 6 oder ein mspD-Gen gemäß Sequenz 8 benutzt
15 wird.

11. Verfahren nach einem der vorhergehenden Ansprüche, wobei
zur Überexpression ein von den Sequenzen 1, 6 oder 8 abgelei-
tetes mutiertes Gen benutzt wird, wobei die Mutation so aus-
20 gebildet ist, daß die chemische und thermische Stabilität sowie die kanalartige Struktur des expimierten Proteins im wesentlichen denen von MspA, MspC oder MspD entspricht.

12. Verfahren nach einem der vorhergehenden Ansprüche, wobei
25 die Mutation im wesentlichen in einer Angleichung der Codons des mspA-Gens, des mspC-Gens oder des mspD-Gens an die Codons der in *E. coli* hoch exprimierten Gene besteht.

13. Verfahren nach einem der vorhergehenden Ansprüche, wobei
30 zur Überexpression ein mutiertes mspA-, mspC- oder mspD-Gen benutzt wird, wobei die Mutation im wesentlichen darin besteht, daß der GC-Gehalt auf weniger als 66% vermindert ist.

14. Verfahren nach einem der vorhergehenden Ansprüche, wobei zur Überexpression das synmspA-Gen gemäß Sequenz 4 benutzt wird.
- 5 15. Verfahren nach einem der vorhergehenden Ansprüche, wobei ein zur Überexpression in *E. coli* geeigneter Vektor verwendet wird, in den das synmspA-Gen gemäß Sequenz 4 eingesetzt ist.
- 10 16. Verfahren nach einem der vorhergehenden Ansprüche, wobei die kanalbildenden Proteine mittels nicht-ionischer oder zwitterionischer Detergentien aus der Zellwand von Gram-positiven Bakterien gewonnen werden.
- 15 17. Verfahren nach einem der vorhergehenden Ansprüche, wobei die Detergentien aus der folgenden Gruppe ausgewählt sind: Isotridecylpoly(ethyleneglycolether)_n, Alkylglucoside, besonders Octylglucosid, Alkylmaltoside, besonders Dodecylmaltosid, Alkylthioglucoside, besonders Octylthioglucosid, Octyl-Polyethylenoxide und Lauryldimethylaminoxid.
- 20 18. Verfahren nach einem der vorhergehenden Ansprüche, wobei die Temperatur bei der Extraktion zwischen 80 und 110°C, vorzugsweise zwischen 90 und 100°C beträgt.
- 25 19. Verfahren nach einem der vorhergehenden Ansprüche, wobei die Extraktionszeit 5 bis 120 Minuten, vorzugsweise 25 - 35 Minuten, beträgt.
- 30 20. Verfahren nach einem der vorhergehenden Ansprüche, wobei ein Puffer mit einer Ionenstärke von mehr als 50 mM NaCl oder Na-Phosphat benutzt wird.

21. Verfahren nach einem der vorhergehenden Ansprüche, wobei das kanalbildende Protein zur Aufreinigung, insbesondere mittels Aceton, ausgefällt wird.
- 5 22. Verfahren nach einem der vorhergehenden Ansprüche, wobei das kanalbildende Protein zur Aufreinigung einer Ionenaustauscher-Chromatographie, insbesondere einer Anionenaustauscher-Chromatographie, unterworfen wird.
- 10 23. Verfahren nach einem der vorhergehenden Ansprüche, wobei das kanalbildende Protein zur Aufreinigung einer Größenausschluß-Chromatographie unterworfen wird.
24. Verfahren nach einem der vorhergehenden Ansprüche, wobei
15 das durch heterologe Überexpression gewonnene kanalbildende Protein durch Erhöhen der lokalen Konzentration des kanalbildenden Proteins renaturiert wird.
25. Verfahren nach Anspruch 24, wobei das Erhöhen der lokalen
20 Konzentration durch elektrophoretische Anreicherung, insbesondere durch Anlegen einer Gleichspannung, durch Ausfällen oder durch Adsorption an einer Oberfläche, insbesondere einer Membran, erfolgt.
- 25 26. Kanalbildendes Protein aus einem Gram-positiven Bakterium hergestellt nach dem Verfahren nach einem der vorhergehenden Ansprüche.
27. Kanalbildendes Protein aus einem Gram-positiven Bakteri-
30 um, wobei das kanalbildende Protein ein Porin ist, das im wesentlichen gegenüber organischen Lösungsmitteln chemisch stabil ist.

28. Kanalbildendes Protein aus einem Gram-positiven Bakterium, wobei das kanalbildende Protein ein Porin ist, das im wesentlichen bis zu einer Temperatur von 80°C thermisch stabil ist.

5

29. Kanalbildendes Protein nach Anspruch 28, wobei das kanalbildende Protein ein Porin ist, das im wesentlichen bis zu einer Temperatur von 100°C thermisch stabil ist.

10 30. Kanalbildendes Protein aus einem Gram-positiven Bakterium, wobei das kanalbildende Protein das Porin MspA, MspC, MspD, ein Fragment dieser Porine, ein zu diesen Porinen oder deren Fragmenten homologes Protein oder ein von der Sequenz dieser Porine abgeleitetes Protein ist.

15

31. Kanalbildendes Protein nach Anspruch 30, wobei die chemische und thermische Stabilität sowie die kanalartige Struktur des abgeleiteten Proteins im wesentlichen derjenigen der Proteine MspA, MspC oder MspD entspricht.

20

32. Gen codierend für ein kanalbildendes Protein nach einem der Ansprüche 26 - 31.

25 33. Gen nach Anspruch 32, wobei das Gen das mspA-Gen gemäß Sequenz 1 ist.

34. Gen nach Anspruch 32, wobei das Gen das mspC-Gen gemäß Sequenz 6 ist.

30 35. Gen nach Anspruch 32, wobei das Gen das mspD-Gen gemäß Sequenz 8 ist.

36. Mutiertes mspA-Gen, mspC-Gen oder mspD-Gen, wobei die Mutation im wesentlichen in einer Angleichung der Codons des mspA-Gens, des mspC-Gens oder des mspD-Gens an die Codons der in *E. coli* hoch exprimierten Gene besteht.

5

37. Mutiertes mspA-Gen, mspC-Gen oder mspD-Gen, insbesondere nach Anspruch 36, wobei die Mutation im wesentlichen darin besteht, daß der GC-Gehalt auf weniger als 66% vermindert ist.

10

38. Mutiertes mspA-Gen, mspC-Gen oder mspD-Gen, insbesondere nach Anspruch 36 oder 37, abgeleitet von einer der Sequenzen 1, 6 oder 8, wobei die Mutation so ausgebildet ist, daß die chemische und thermische Stabilität sowie die kanalartige Struktur des exprimierten Proteins im wesentlichen der von MspA, MspC oder MspD entspricht.

15

39. Mutiertes mspA-Gen nach einem der Ansprüche 36 bis 38, wobei das mutierte Gen das synmspA-Gen gemäß Sequenz 4 ist.

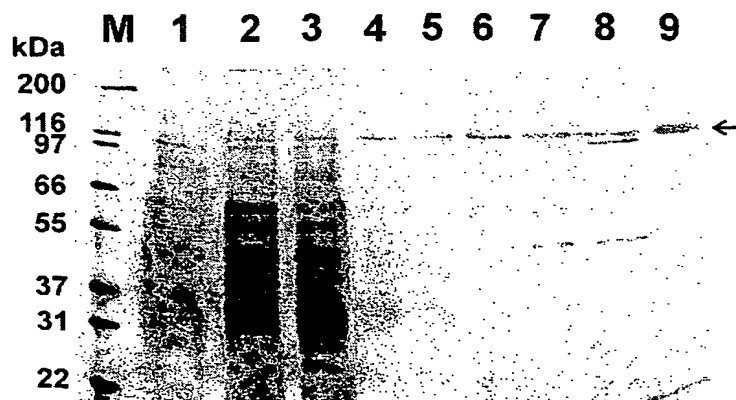
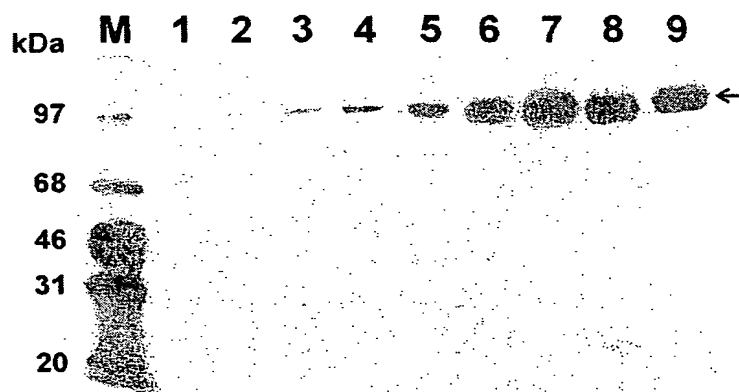
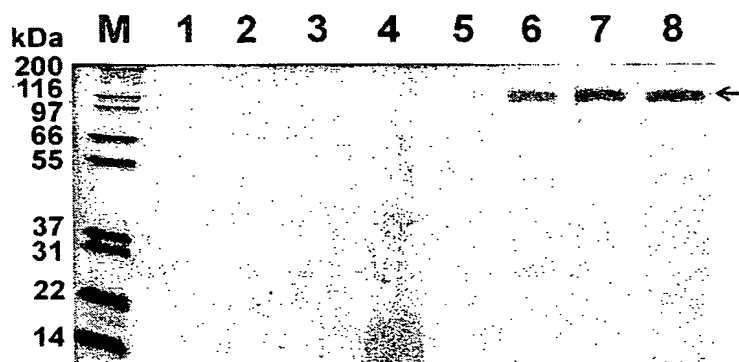
20

40. Plasmidvektor pMN501.

41. Überexpressionssystem, bei dem *E. coli* den Plasmidvektor pMN501 enthält.

25

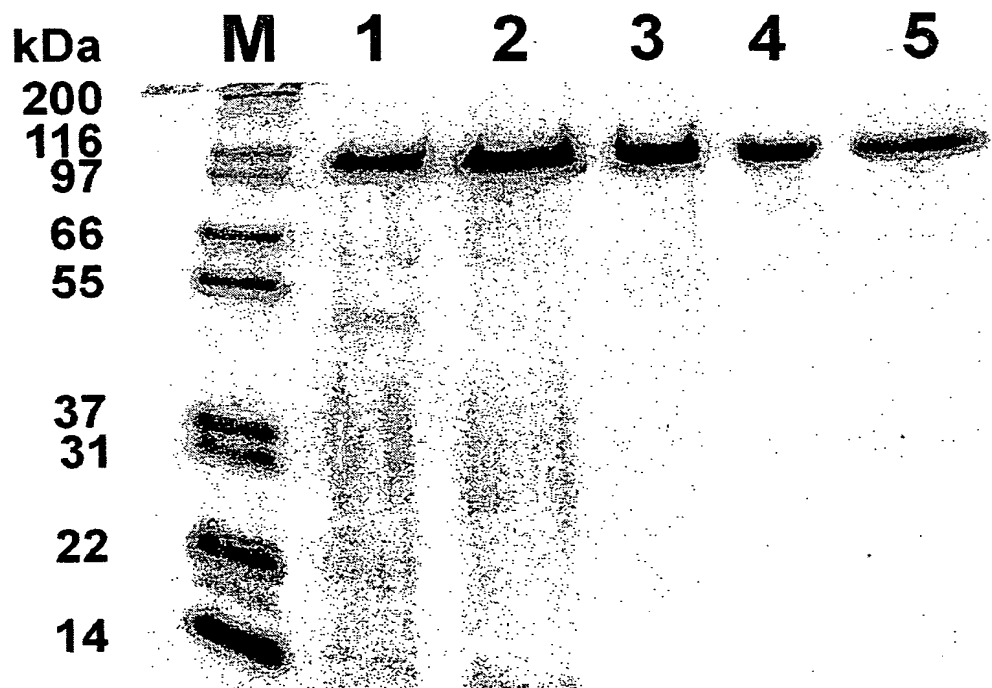
1/6



BEST AVAILABLE COPY

This Page Blank (uspto)

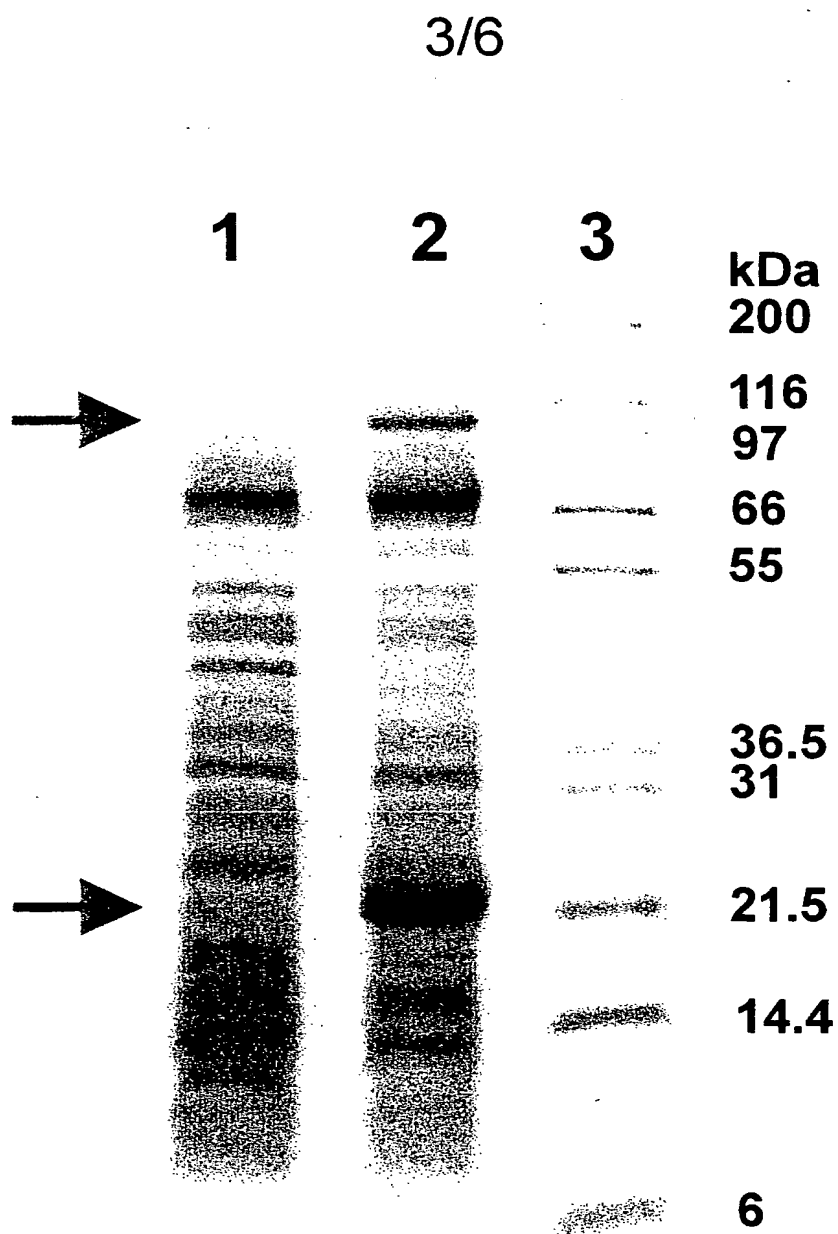
2/6



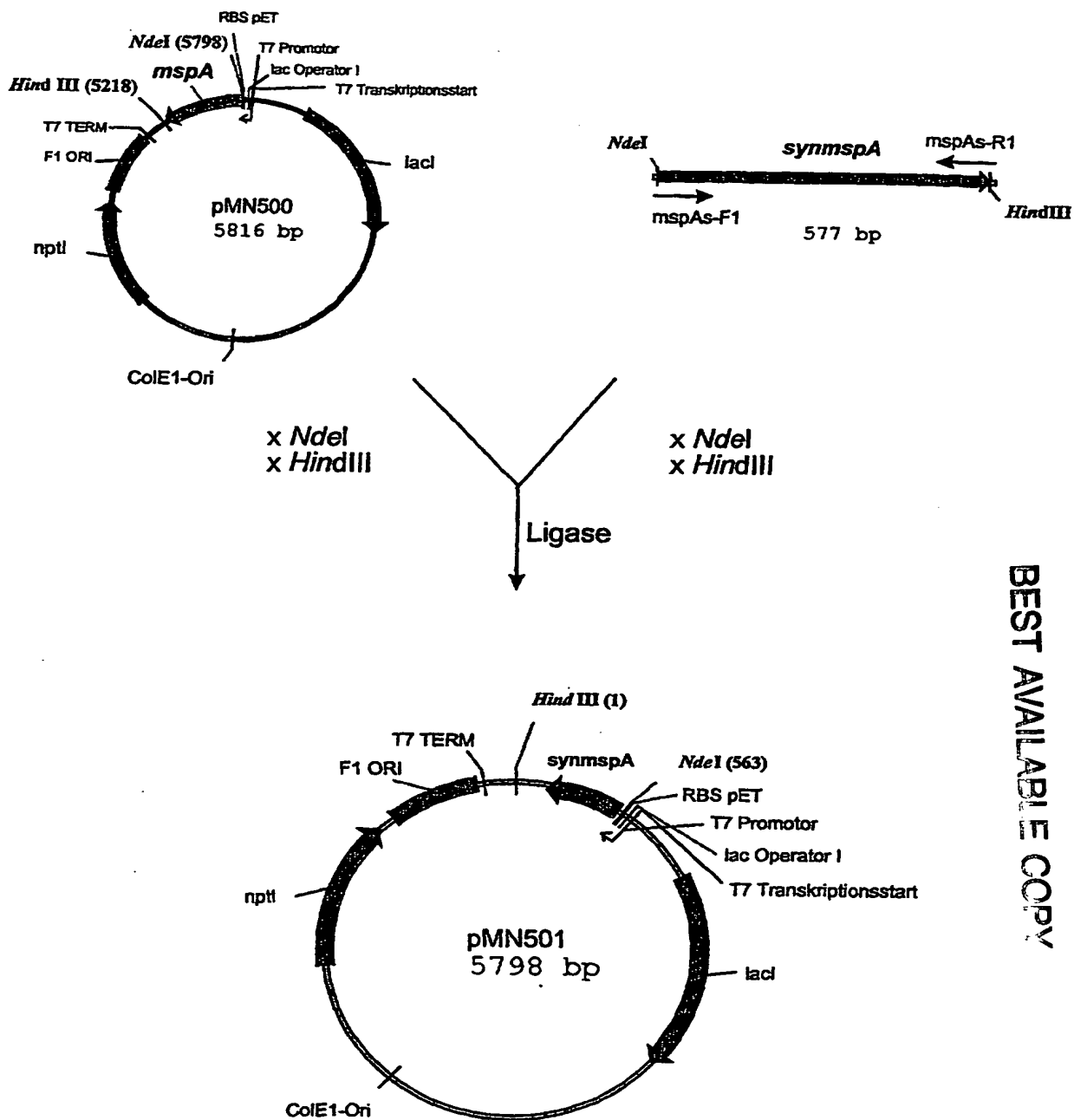
BEST AVAILABLE COPY

Fig. 2

This Page Blank (uspto)

*Fig. 3*

This Page Blank (uspto)



BEST AVAILABLE COPY

Fig. 4

This Page Blank (uspto)

5/6

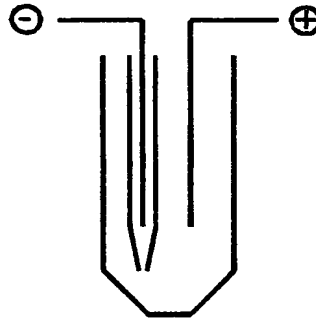


Fig. 5

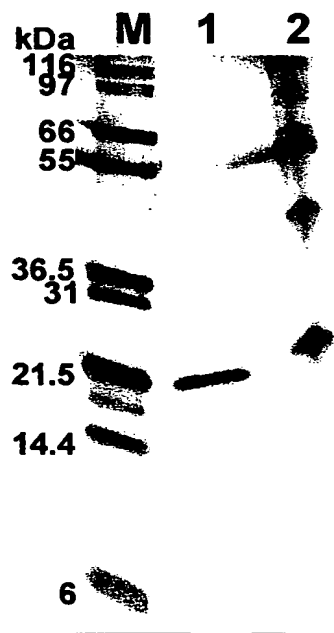


Fig. 6

BEST AVAILABLE COPY

This Page Blank (uspio)

6/6

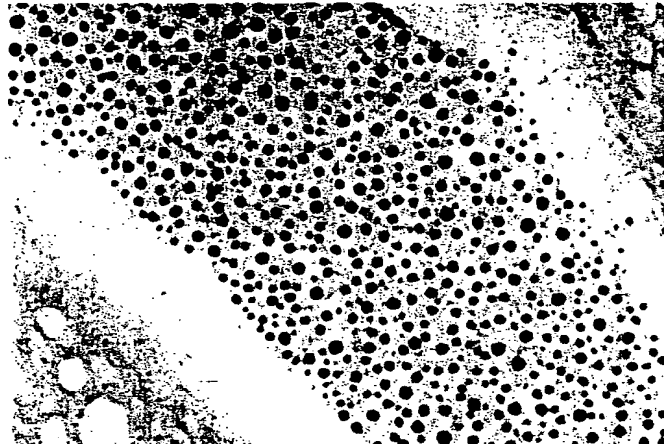


Fig. 7a

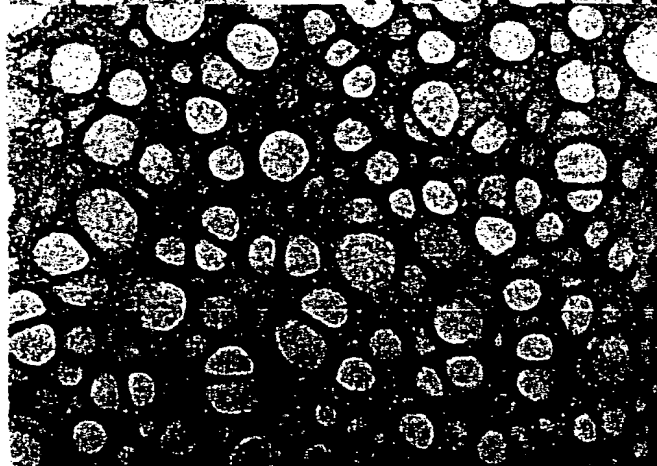


Fig. 7b

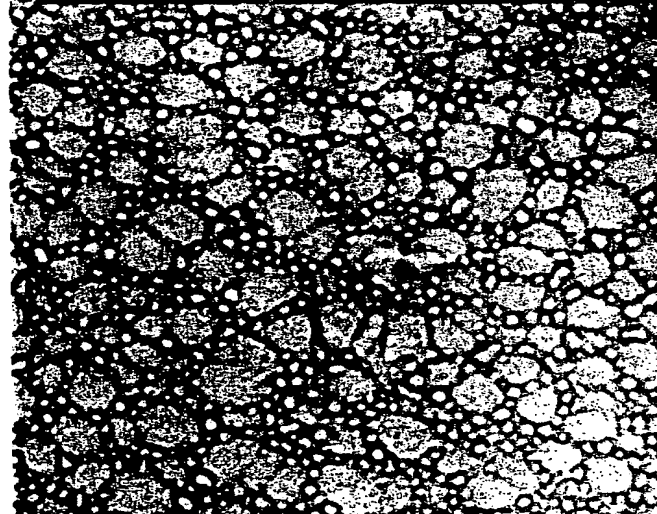


Fig. 7c

BEST AVAILABLE COPY

This Page Blank (uspto)

SEQUENZPROTOKOLL

5 <110> Niederweis Dr., Michael
 Bossmann Dr., Stefan
 <120> Verfahren zur Herstellung eines kanalbildenden Proteins
 <130> 401172
 10 <140>
 <141>
 <150> DE 199 43 520.0
 15 <151> 1999-09-11
 <150> DE 199 41 416.5
 <151> 1999-08-31
 20 <160> 9
 <170> PatentIn Ver. 2.1
 <210> 1
 25 <211> 636
 <212> DNA
 <213> Mycobacterium smegmatis
 <220>
 30 <221> CDS
 <222> (1) .. (636)
 <223> mspA-Gen
 <400> 1
 35 atg aag gca atc agt cgg gtg ctg atc gcg atg gtt gca gcc atc gcg 48
 Met Lys Ala Ile Ser Arg Val Leu Ile Ala Met Val Ala Ala Ile Ala
 1 5 10 15
 40 gcg ctt ttc acg agc aca ggc acc tct cac gca ggc ctg gac aac gag 96
 Ala Leu Phe Thr Ser Thr Gly Thr Ser His Ala Gly Leu Asp Asn Glu
 20 25 30
 45 ctg agc ctc gtt gat ggc cag gac cgc acc ctc acc gtg cag cag tgg 144
 Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu Thr Val Gln Gln Trp
 35 40 45
 50 gac acc ttc ctc aat ggt gtg ttc ccc ctg gac cgc aac cgt ctt acc 192
 Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp Arg Asn Arg Leu Thr
 50 55 60
 55 cgt gag tgg ttc cac tcc ggt cgc gcc aag tac atc gtg gcc ggc ccc 240
 Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr Ile Val Ala Gly Pro
 65 70 75 80
 60 ggt gcc gac gag ttc gag ggc acg ctg gaa ctc ggc tac cag atc ggc 288
 Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu Gly Tyr Gln Ile Gly
 85 90 95
 ttc ccg tgg tcg ctg ggt gtg ggc atc aac ttc agc tac acc acc ccg 336
 Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe Ser Tyr Thr Thr Pro
 100 105 110

This Page Blank (uspto)


```

5   aac atc ctg atc gac gac ggt gac atc acc gct ccg ccg ttc ggc ctg 384
    Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr Ala Pro Pro Phe Gly Leu
      115                                120                                125

10  gat ctg ggc aac ggc ccc ggc atc cag gaa gtc gca acg ttc tcg gtc 480
    Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val Ala Thr Phe Ser Val
    145                                150                                155                                160

15  gac gtc tcc ggc gcc gag ggt ggc gtg gcc gtg tcg aac gcc cac ggc 528
    Asp Val Ser Gly Ala Glu Gly Gly Val Ala Val Ser Asn Ala His Gly
      165                                170                                175

20  acc gtg acc ggt gcg gcc ggc ggt gtg ctg ctg cgt ccg ttc gcc cgc 576
    Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu Arg Pro Phe Ala Arg
      180                                185                                190

25  ctg atc gcc tcg acc ggt gac tcg gtc acc acc tac ggc gaa ccc tgg 624
    Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr Tyr Gly Glu Pro Trp
      195                                200                                205

30  aac atg aac tga 636
    Asn Met Asn
      210

35  <210> 2
    <211> 1423
    <212> DNA
    <213> Mycobacterium smegmatis

40  <220>
    <221> -10_signal
    <222> (323)..(328)
    <223> vermuteter Promotor

45  <220>
    <221> CDS
    <222> (499)..(1134)
    <223> mspA-Gen

50  <220>
    <221> RBS
    <222> (492)..(496)
    <223> vermutete Ribosomenbindestelle

55  <400> 2
    gttaacggag tcgggccgtc gatacggcgg cgaagatcat ccggcagatt ggcgccctggt 60
    taaacccgcg taaacactgg taccgccggt ccgcgccgga aaaggttttg cctcacgggtg 120
    aatatgtgac ctgaattgca cttcacgggt aaaagcggag gtaaccgacg gttgccgcag 180
60  caccctcaca gcttgggcca aggtgacgtg cagcgcacgc ctgccggtgc cggatggcgg 240

```

This Page Blank (uspto)

	tcaccgcaaa	gtgtcaggca	ctgccgaaaag	gtcagtcagc	aaacttcact	gcggtctgtg	300
	tgcgaaagtgc	ggttgtgga	cgtatccgtt	gctgccgcgc	gccctggcgt	ttatgtttct	360
5	gctgccaaact	gtgagcgagg	cattagagac	agatgtgac	ctcttagatc	tccgaagtct	420
	ctgaacaggt	gttgagccgg	ttgcagacaa	caaaacaggt	gggcctgagg	ggccgccggc	480
10	gatacagtta	gggagaac	atg aag gca atc agt cgg gtg ctg atc gcg atg	531			
		Met Lys Ala Ile Ser Arg Val Leu Ile Ala Met					
		1		5		10	
	gtt gca gcc atc gcg gcg ctt ttc acg agc aca ggc acc tct cac gca	579					
15	Val Ala Ala Ile Ala Ala Leu Phe Thr Ser Thr Gly Thr Ser His Ala						
		15		20		25	
	ggc ctg gac aac gag ctg agc ctc gtt gat ggc cag gac cgc acc ctc	627					
20	Gly Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu						
		30		35		40	
	acc gtg cag cag tgg gac acc ttc ctc aat ggt gtg ttc ccc ctg gac	675					
	Thr Val Gln Gln Trp Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp						
		45		50		55	
25	cgc aac cgt ctt acc cgt gag tgg ttc cac tcc ggt cgc gcc aag tac	723					
	Arg Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr						
		60		65		70	
	atc gtg gcc ggc ccc ggt gcc gac gag ttc gag ggc acg ctg gaa ctc	771					
30	Ile Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu						
		80		85		90	
	ggc tac cag atc ggc ttc ccg tgg tcg ctg ggt gtg ggc atc aac ttc	819					
35	Gly Tyr Gln Ile Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe						
		95		100		105	
	agc tac acc acc ccg aac atc ctg atc gac gac ggt gac atc acc gct	867					
40	Ser Tyr Thr Thr Pro Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr Ala						
		110		115		120	
	ccg ccg ttc ggc ctg aac tcg gtc atc acc ccg aac ctg ttc ccc ggt	915					
	Pro Pro Phe Gly Leu Asn Ser Val Ile Thr Pro Asn Leu Phe Pro Gly						
		125		130		135	
45	gtg tcg atc tcg gca gat ctg ggc aac ggc ccc ggc atc cag gaa gtc	963					
	Val Ser Ile Ser Ala Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val						
		140		145		150	
	gca acg ttc tcg gtc gac gtc tcc ggc gcc gag ggt ggc gtg gcc gtg	1011					
50	Ala Thr Phe Ser Val Asp Val Ser Gly Ala Glu Gly Gly Val Ala Val						
		160		165		170	
	tcg aac gcc cac ggc acc gtg acc ggt ggc gcc ggc ggt gtg ctg ctg	1059					
55	Ser Asn Ala His Gly Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu						
		175		180		185	
	cgt ccg ttc gcc cgc ctg atc gcc tcg acc ggt gac tcg gtc acc acc	1107					
60	Arg Pro Phe Ala Arg Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr						
		190		195		200	
	tac ggc gaa ccc tgg aac atg aac tga ttcttgacc gccgttcggt	1154					

This Page Blank (uspto)

Tyr Gly Glu Pro Trp Asn Met Asn
 205 210

5 cgctgagacc gcttgagatc ggcgcgtccc gctcccgggtg tcgtcagetc atcgttgaca 1214
 cgtgaactga cactcttcct agccggagcg kacgcgccga tcttgtgttc tgagcagttc 1274
 tcagtccgtc cgccgcaaca ccagcgtga cggcgtacgc agcctgcccc ccaccgcgcg 1334
 10 ccagggacgc cccagcctgg gcaccacctc agcggtcggc acgatgcgcg gatcgggtcac 1394
 ctggaacgtc tcaccgttca tcaccgcgcg 1423

15

<210> 3
 <211> 211
 <212> PRT
 20 <213> Mycobacterium smegmatis

25 <220>
 <221> SIGNAL
 <222> (1)..(27)
 <223> vermutete Signalsequenz des MspA-Proteins

30 <220>
 <221> PEPTIDE
 <222> (28)..(211)
 <223> reifes MspA-Protein

35 <400> 3
 Met Lys Ala Ile Ser Arg Val Leu Ile Ala Met Val Ala Ala Ile Ala
 1 5 10 15
 Ala Leu Phe Thr Ser Thr Gly Thr Ser His Ala Gly Leu Asp Asn Glu
 20 25 30
 40 Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu Thr Val Gln Gln Trp
 35 40 45
 Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp Arg Asn Arg Leu Thr
 50 55 60
 Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr Ile Val Ala Gly Pro
 65 70 75 80
 50 Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu Gly Tyr Gln Ile Gly
 85 90 95
 Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe Ser Tyr Thr Thr Pro
 100 105 110
 55 Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr Ala Pro Pro Phe Gly Leu
 115 120 125
 Asn Ser Val Ile Thr Pro Asn Leu Phe Pro Gly Val Ser Ile Ser Ala
 60 130 135 140

This Page Blank (uspto)

Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val Ala Thr Phe Ser Val
 145 150 155 160
 5 Asp Val Ser Gly Ala Glu Gly Gly Val Ala Val Ser Asn Ala His Gly
 165 170 175
 Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu Arg Pro Phe Ala Arg
 180 185 190
 10 Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr Tyr Gly Glu Pro Trp
 195 200 205
 Asn Met Asn
 210
 15
 20 <210> 4
 <211> 558
 <212> DNA
 <213> Künstliche Sequenz
 25 <220>
 <223> Beschreibung der künstlichen Sequenz: synthetisch
 <220>
 <221> CDS
 30 <222> (1)..(558)
 <223> synmspA-Gen
 <400> 4
 35 atg ggc ctg gac aac gaa ctg tcc ctg gtt gac ggc cag gac cgt acc 48
 Met Gly Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr
 1 5 10 15
 ctg acc gtt cag cag tgg gac acc ttc ctg aac ggt gtt ttc ccg ctg 96
 40 Leu Thr Val Gln Gln Trp Asp Thr Phe Leu Asn Gly Val Phe Pro Leu
 20 25 30
 gac cgt aac cgt ctg acc cgt gaa tgg ttc cac tcc ggt cgt gcg aaa 144
 45 Asp Arg Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Lys
 35 40 45
 tac atc gtt gcg ggt ccg ggt gcg gac gag ttc gaa ggt acc ctg gaa 192
 Tyr Ile Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu
 50 55 60
 50 ctg ggt tac cag atc ggc ttc ccg tgg tcc ctg ggt gtt ggt atc aac 240
 Leu Gly Tyr Gln Ile Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn
 65 70 75 80
 ttc tct tac acc acc ccg aac atc ctg atc gac gac ggt gac atc acc 288
 55 Phe Ser Tyr Thr Thr Pro Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr
 85 90 95
 gct ccg ccg ttc ggt ctg aac tct gtt atc acc ccg aac ctg ttc ccg 336
 60 Ala Pro Pro Phe Gly Leu Asn Ser Val Ile Thr Pro Asn Leu Phe Pro
 100 105 110

This Page Blank (uspto)

ggt gtt tct atc tct gct gat ctg ggc aac ggt ccg ggt atc cag gaa 384
 Gly Val Ser Ile Ser Ala Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu
 115 120 125

5 gtt gct acc ttc tct gta gac gtc tct ggt gct gaa ggt ggt gtt gct 432
 Val Ala Thr Phe Ser Val Asp Val Ser Gly Ala Glu Gly Gly Val Ala
 130 135 140

10 gtt tct aac gct cac ggc acc gtt acc ggt gcg gct ggc ggt gtt ctg 480
 Val Ser Asn Ala His Gly Thr Val Thr Gly Ala Ala Gly Gly Val Leu
 145 150 155 160

15 ctg cgt ccg ttc gct cgt ctg atc gct tct acc ggt gac tct gtt acc 528
 Leu Arg Pro Phe Ala Arg Leu Ile Ala Ser Thr Gly Asp Ser Val Thr
 165 170 175

acc tac ggt gaa ccg tgg aac atg aac tga 558
 Thr Tyr Gly Glu Pro Trp Asn Met Asn
 180 185

20

<210> 5
 25 <211> 185
 <212> PRT
 <213> Künstliche Sequenz

<220>
 30 <221> PEPTIDE
 <222> (1)..(184)
 <223> rMspA

35 <220>
 <223> Beschreibung der künstlichen Sequenz:synthetisch

<400> 5
 40 Met Gly Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr
 1 5 10 15
 Leu Thr Val Gln Gln Trp Asp Thr Phe Leu Asn Gly Val Phe Pro Leu
 20 25 30

45 Asp Arg Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Lys
 35 40 45
 Tyr Ile Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu
 50 50 55 60

50 Leu Gly Tyr Gln Ile Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn
 65 70 75 80

55 Phe Ser Tyr Thr Thr Pro Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr
 85 90 95

Ala Pro Pro Phe Gly Leu Asn Ser Val Ile Thr Pro Asn Leu Phe Pro
 100 105 110

60 Gly Val Ser Ile Ser Ala Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu
 115 120 125

This Page Blank (uspto)

Val Ala Thr Phe Ser Val Asp Val Ser Gly Ala Glu Gly Gly Val Ala
 130 135 140

5 Val Ser Asn Ala His Gly Thr Val Thr Gly Ala Ala Gly Gly Val Leu
 145 150 155 160

Leu Arg Pro Phe Ala Arg Leu Ile Ala Ser Thr Gly Asp Ser Val Thr
 165 170 175

10 Thr Tyr Gly Glu Pro Trp Asn Met Asn
 180 185

15

<210> 6
 <211> 648
 <212> DNA
 20 <213> Mycobacterium smegmatis

<400> 6
 atgaaggcaa tcagtcgggt gctgatcgcg atgatttccg cgttggtcgc ggccgctcgc 60
 gggttgttcg tgagcgcggg caccctctcac gcgggtctcg acaatgagct cagccttgtc 120
 25 gatggtcagg accgcaccct caccgtgcag cagtgggata cgttcctcaa tgggtgtgttc 180
 cccctggacc gcaaccgtct gaccctgag tggttccact ccggtcgcgc gaagtacatc 240
 gtggccggcc ccggtgccga tgagttcgag ggcacgctgg aactcggcta ccagatcggc 300
 ttcccgtggt cgctgggtgt gggcatcaac ttcagctaca ccaccccgaa catcctgac 360
 gacgacgggt acatcaccgg tccgcccttc ggccctcgagt cggtcacac cccgaacctg 420
 30 ttccccgggt tgatcgtctc ggccgacctg ggcaacggcc ccggcatcca ggaagtgcgc 480
 acgttctcgg tcgacgtctc gggctccgca ggccggagtag cggctctcaa cgcgcacggc 540
 accgtcaccg gtgcggccgg cggtgtgctg ctgcgtccgt tcgcccgcct gatcgccctg 600
 accggtgact cggtcaccac ctacggcgaa ccctggaaca tgaactga 648

35

<210> 7
 <211> 184
 40 <212> PRT
 <213> Mycobacterium smegmatis

<400> 7
 Gly Leu Asp Asn Glu Leu Ser Leu Val Asp Gly Gln Asp Arg Thr Leu
 45 1 5 10 15

Thr Val Gln Gln Trp Asp Thr Phe Leu Asn Gly Val Phe Pro Leu Asp
 20 25 30

50 Arg Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Lys Tyr
 35 40 45

Ile Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu
 50 55 60

55 Gly Tyr Gln Ile Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe
 65 70 75 80

Ser Tyr Thr Thr Pro Asn Ile Leu Ile Asp Asp Gly Asp Ile Thr Gly
 60 85 90 95

This Page Blank (uspto)

Pro Pro Phe Gly Leu Glu Ser Val Ile Thr Pro Asn Leu Phe Pro Gly
 100 105 110
 5 Val Ser Ile Ser Ala Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val
 115 120 125
 Ala Thr Phe Ser Val Asp Val Ser Gly Pro Ala Gly Gly Val Ala Val
 130 135 140
 10 Ser Asn Ala His Gly Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu
 145 150 155 160
 Arg Pro Phe Ala Arg Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr
 165 170 175
 15 Tyr Gly Glu Pro Trp Asn Met Asn
 180
 20
 <210> 8
 <211> 624
 <212> DNA
 <213> Mycobacterium smegmatis
 25
 <400> 8
 gtgcgctacc tcgtcatgat gttcgtctcta ctctgtgtccg tgacgcttgt gagccccgc 60
 cctgccaaacg cgggtggacaa tcagctcagc gtggtcgcacg gccaaaggctc cacgctgacc 120
 30 gtgcagcaag ccgagacatt cctcaacggc gtgttccttc tcgaccggaa ccgactgacc 180
 cgtgagtggt ttactccgg ccgcgccacc taccatgtgg ccggcccagg tgccgacgaa 240
 ttcgagggca cgctcgaact cgggtatcag gtcggcttcc cgtggtcatt gggcgtcggc 300
 atcaacttct cgtacacgac cccgaacatc ctcatcgacg gaggcgacat caccagccg 360
 ccgttcggcc tggacacccat catcaccccc aacctcttcc ccggcgtgtc catcagtgcc 420
 35 gacctcggca acggtcccg taccaggag gtcgccacct tctcgggtgga cgtgaagggc 480
 gcgaaaggag cggtcgccgt atccaatgcg catggcaccg tgaccggcgc ggccggcggc 540
 gtgctcctgc gtccgttcgc ccggttgatc gcctcgacgg gcgacagcgt caccacctac 600
 ggcgagccct ggaacatgaa ctag 624
 40
 <210> 9
 <211> 183
 <212> PRT
 45 <213> Mycobacterium smegmatis
 <400> 9
 Val Asp Asn Gln Leu Ser Val Val Asp Gly Gln Gly Arg Thr Leu Thr
 1 5 10 15
 50 Val Gln Gln Ala Glu Thr Phe Leu Asn Gly Val Phe Pro Leu Asp Arg
 20 25 30
 55 Asn Arg Leu Thr Arg Glu Trp Phe His Ser Gly Arg Ala Thr Tyr His
 35 40 45
 Val Ala Gly Pro Gly Ala Asp Glu Phe Glu Gly Thr Leu Glu Leu Gly
 50 55 60
 60 Tyr Gln Val Gly Phe Pro Trp Ser Leu Gly Val Gly Ile Asn Phe Ser
 65 70 75 80

This Page Blank (uspto)

Tyr Thr Thr Pro Asn Ile Leu Ile Asp Gly Gly Asp Ile Thr Gln Pro
85 90 95

5 Pro Phe Gly Leu Asp Thr Ile Ile Thr Pro Asn Leu Phe Pro Gly Val
100 105 110

Ser Ile Ser Ala Asp Leu Gly Asn Gly Pro Gly Ile Gln Glu Val Ala
115 120 125

10 Thr Phe Ser Val Asp Val Lys Gly Ala Lys Gly Ala Val Ala Val Ser
130 135 140

15 Asn Ala His Gly Thr Val Thr Gly Ala Ala Gly Gly Val Leu Leu Arg
145 150 155 160

Pro Phe Ala Arg Leu Ile Ala Ser Thr Gly Asp Ser Val Thr Thr Tyr
165 170 175

20 Gly Glu Pro Trp Asn Met Asn
180

This Page Blank (uspto)